Journal of Phonetics 78 (2020) 100934

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Special Issue: Integrating Phonetics and Phonology, eds. Cangemi & Baumann

Dialectal phonology constrains the phonetics of prominence

Rachel Smith [a,*], Tamara Rathcke [b]

[a] University of Glasgow, United Kingdom
[b] University of Kent, United Kingdom

A R T I C L E I N F O

Article history:
Received 6 April 2018
Received in revised form 25 July 2019
Accepted 3 October 2019
Available online 9 December 2019

Keywords:
Prominence
Prominence gradient
Phonetic cues
Timing
Glasgow english
Standard Southern British English
Dialectal differences

A B S T R A C T

Accentual prominence has well-documented effects on various phonetic properties, including timing, vowel quality, amplitude, and pitch. These cues can exist in trading relationships and can differ in magnitude in different languages. Less is understood about how phonetic cues to accentuation surface under different phonological constraints, such as those posed by segmental phonology, aspects of the prosodic hierarchy, and intonational phonology. Dialectal comparisons offer a valuable window on these issues, because dialects of a language share basic aspects of structure and function, but can differ in key segmental and suprasegmental constraints which may affect the cues that realise accentual prominence. We compared the realisation of trochaic words (e.g. *cheesy, picky*) in accented/unaccented and phrase-final/non-final positions in two dialects of British English, Standard Southern British English, and Standard Scottish English as spoken in Glasgow. We found generally shallower prominence gradients for Glasgow than SSBE with respect to intensity and duration, and very little evidence of accentual lengthening of vowels in Glasgow, compared to robust effects in SSBE. In contrast, phrase-finality had similar effects across the two dialects. The differences observed illustrate how the expression of accentual prominence reflects and reveals the different segmental and intonational systems that operate within dialects of the same language.

© 2019 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Parts of the speech stream stand out when someone speaks: those elements are prominent. Prominence has many phonetic ingredients, including both traditionally "prosodic" and traditionally "segmental" parameters. This study explores whether and how these ingredients combine differently in different dialects of English, according to their specific segmental and intonational phonological constraints. The answer to this question may help to understand the balance of universal and specific factors in how prominence is controlled, and how prominence is perceived despite variation in production.

Prominence is a broad term, interpreted differently by different researchers (Gussenhoven, 2011, 2015). Perceptually, a range of factors conspire to determine what stands out in the ebb and flow of speech, from intrinsic acoustic salience through to familiarity, predictability, attentional demands and the type of task the listener is engaged in (Cole, Mo, & Hasegawa-Johnson, 2010; Cole, Mo, & Baek, 2010; Baumann & Winter, 2018; Bishop, Kuo, & Kim, 2019; Cole et al., 2019). In this sense, prominence encompasses more than the linguistic structure of an utterance. Nonetheless, from a linguistic standpoint, for English specifically, the prosodic systems of stress and (intonational) accentuation are what we mainly refer to when discussing prominence. For some authors, English stress and accent represent different degrees of prominence, which are formalisable as levels in a prosodic hierarchy that is organised around metrical principles (e.g. Calhoun, 2010; Ladd, 2008). Others disagree that phonological prominence is a unified phenomenon of which stress and accentuation express different degrees, and instead see them as drastically different linguistic systems: word-stress is representable in terms of metrical feet, while sentence prosody is representable in terms of rules governing phonological phrasing and pitch accent distribution (Gussenhoven, 2011, 2015). We do not take a strong view on this issue in this paper; instead, our focus is on the phonetic correlates of accentuation (henceforth *accentual prominence*), and the extent to which these are dialect-specific. We also consider

* Corresponding author.
E-mail addresses: rachel.smith@glasgow.ac.uk (R. Smith), t.v.rathcke@kent.ac.uk (T. Rathcke).

https://doi.org/10.1016/j.wocn.2019.100934
0095-4470/© 2019 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the dialect-specific correlates of position in phrase, with specific attention to finality.

1.1. Phonetic cues to accentual prominence

Accentual prominence can be reflected in many phonetic parameters, including the size and shape of pitch excursions, amplitude, duration, spectral tilt and other features of the voice source, vowel quality, segmental hyperarticulation (including magnitude of opening and closing gestures and resistance to coarticulation), and even multi-modal co-speech gestures (e.g. Cho, 2004; de Jong, 1995; Terken, 1991, 1994; Turk & Sawusch, 1996; Wagner et al., 2015).

In theory, accentual prominence could be a matter of "turning up the gain" on these parameters: producing larger pitch excursions, greater amplitude, more lengthening, and longer and stronger gestures. Correspondingly, finality might be a matter of "winding down the clock": slowing the duration of gestures, reducing amplitude, and other features of "supra-laryngeal declination" (Vayra & Fowler, 1992). If this were the case, we would not expect the phonetics of these aspects of prosodic structure to vary much cross-dialectally. However, the reality is more complex, for two reasons. First, every parameter that contributes to accentuation plays a part in other linguistic functions as well, and these functions impose their own constraints. Second, the phonetic parameters that are involved in accentuation may not be independent of one another physiologically or in terms of control mechanisms, so trading relations can occur.

There is already some evidence that the way accentual prominence is achieved depends on the system of linguistic contrasts operating within a language. An elegant demonstration of this was made by de Jong (1995), who investigated whether the differences between nuclear accented, prenuclear accented and stressed but unaccented syllables are cued by overall articulatory parameters such as increased jaw opening or increased sonority (cf. Smith, Erickson, & Savariaux, 2019), or rather whether the expression of these differences depends on the system of linguistic contrasts operating within a language. Using X-ray microbeam data, he demonstrated that simple sonority or jaw opening scales could not explain the patterns of variation in American English, where low vowels achieved a lower tongue/jaw position under accentuation, high back vowels a backer position, and front vowels a higher and fronter position. De Jong concluded that accentuation involves *localized shifts towards hyperarticulated speech,* and further, that hyperarticulation patterns under accentuation "can act as a diagnostic for determining the content of the linguistic code of a particular language" (1995: 502), i.e. for determining which phonetic dimensions hold particular linguistic importance in a system. Further research has confirmed this concept with a wider range of languages, e.g. de Jong and Zawaydeh (2002), Cho and McQueen (2005), Oh and Byrd (2019) show for Korean that the realisation of focal prominence in Korean stop-initial syllables is sensitive to the tonal and durational characteristics associated with the stop type (fortis, lenis or aspirated). Meanwhile, other evidence underscores that cues to accentual prominence operate flexibly, exist in trading relationships, and can differ in magnitude in different languages (see Fletcher, 2010 for a review). These issues are explored more fully in Sections 1.2 and 1.3, and the reasons to conduct a cross-dialectal investigation are outlined in Section 1.4.

1.2. Relationships among linguistic functions: The role of timing

Timing is a dimension where the interplay of accentual prominence, phrase-finality and other linguistic functions can be usefully explored. Timing marks many segmental contrasts (e.g. singleton-geminate consonants, vowel distinctions in both quantity and non-quantity languages), as well as accentuation (Turk & Sawusch, 1997; Turk & White, 1999) and grouping at the lexical and phrasal levels (Turk & Shattuck-Hufnagel, 2000, 2007). The timing reflexes of accentuation and phrase-finality have been well-studied in themselves and to some extent in terms of their interaction with each other and with segmental timing phenomena.

Compared to unaccented syllables, accented syllables lengthen, by around 20% for American English according to Turk and Sawusch (1997), and in many other languages studied (e.g. Dutch, Eefting, 1991; Swedish, Heldner & Strangert, 2001; Polish, Oliver & Grice, 2003) though not all (e.g. no accentual lengthening was found for Arabic by de Jong & Zawaydeh, 2002). The effects of accentual prominence on timing may depend on the function of the accent, e.g. broad vs. narrow focus, or contrastive vs. non-contrastive accent (Baumann, Becker, Grice, & Mücke, 2007). Following the line of argument presented by de Jong (1995), various phonological contrasts have been shown to be more clearly marked on accented than non-accented words, e.g. VOT, vowel quantity, intrinsic vowel duration differences, and vowel differences due to following voicing (see Fletcher, 2010 for a review). Research into the domain of cues to accentual prominence has investigated the extent to which these are concentrated and localised upon the accent-bearing unit, distributed within it, and/or able to spread beyond it. Most work suggests that lengthening is most noticeable on the vowel of an accented word, but syllable-initial and -final consonants can be affected too. Moreover, accentual lengthening can spread rightwards to other syllables within a prosodic word (e.g. Turk & White, 1999, Turk & Shattuck-Hufnagel, 2000, Cho & McQueen, 2005, Cho & Keating, 2009) and to a lesser extent leftwards (Cambier-Langeveld & Turk, 1999; Turk & White, 1999).

Syllables that are final in a phrase or larger prosodic domain also lengthen, compared to non-final syllables. Phrase-final lengthening has been observed in a wide range of languages, with no exceptions cited in Fletcher (2010) comprehensive review. The lengthening may be by up to 90% for US English (Turk & Shattuck-Hufnagel, 2007). It can involve spatially larger as well as longer articulatory gestures, and in some cases, slower closing gestures (Fletcher, 2010: 541). These patterns have been interpreted in terms of a general relaxation of speech gestures, akin to declination at the supra-laryngeal level (Lindblom, 1968; Tabain, 2003; Vayra & Fowler, 1992), or as a slowing of the clock that controls activation of speech gestures (Byrd & Saltzman, 2003). Phrase-final lengthening may be cumulative, increasing with the strength of the following boundary, though the evidence is mixed (Klatt, 1975; Umeda, 1975; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). It can begin before the phrase-final syllable; its extent may reflect structural factors, local phonetic factors, or a balance of both (Turk & Shattuck-Hufnagel, 2007) as discussed further below.

Where timing is implicated in one type of function, it may be less available for another. In quantity languages, timing is heavily involved in segmental contrasts, and is less available for marking the prosodic hierarchy. Remijsen and Gilley (2008) for Dinka, Nakai et al. (2012) for Finnish and White and Mády (2008) for Hungarian, all showed that segmental contrast was preserved in final position and, consequently, final lengthening occurred on short vowels only to a limited degree. Conversely, Myers and Hansen (2007) review the way that for Hawaiian, Lithuanian and Tagalog, phonemic contrasts are neutralised towards the short phoneme in phrase- or utterance-final positions, indicating that phrase-final lengthening "wins" over segmental timing in these languages. There has been little investigation of how prosodic and segmental constraints interact in non-quantity languages where duration may be one exponent of vowel identity. Cooper and Danly (1981) found that phrase-final lengthening was significantly greater for vowels that are long by virtue of preceding a voiced consonant, than for vowels that are short because they precede a voiceless consonant (with a preceding short vowel). Turk and Shattuck-Hufnagel (2007) hypothesized that English lax or high vowels might have limited expandability in word-final syllables, with the consequence that phrase-final lengthening might be initiated on earlier syllables in words containing these vowels, but failed to find clear support for this hypothesis in American English.

Lengthening that is associated with prominence interacts with lengthening due to an upcoming prosodic boundary. Turk and Shattuck-Hufnagel (2007) showed for American English that boundary-related lengthening does not only affect the very last syllable in a phrase; if the phrase-final word has stress on an earlier syllable, phrase-final lengthening can begin at the stressed syllable, and can "skip" intermediate regions. They therefore proposed that final lengthening has multiple targets. Byrd and Riggs (2008) obtained similar results in an articulatory investigation, and couched their interpretation in terms of coupling between different types of prosodic gestures in the Articulatory Phonology framework. In articulatory investigations of Greek, Katsika, Krivokapić, Mooshammer, Tiede, and Goldstein (2014) and Katsika (2016) showed that in words with final stress, phrase-final lengthening affects only the articulatory gestures of the phrase-final syllable immediately adjacent to the boundary, but in words with non-final stress, phrase-final lengthening is initiated earlier. Kim, Jang, and Cho (2017) investigated pre-boundary lengthening in trisyllabic pseudo-words (bábaba, babába, bababá) in American English. They demonstrated that pre-boundary lengthening was modulated by the degree of prominence, i.e., the less prominent the pre-boundary syllable, the more pre-boundary lengthening was observed. Pre-boundary lengthening was attracted to the penultimate stressed syllable when the word received no pitch accent, but not otherwise. These findings might suggest a limit on expandability, in parallel to the limits on compressibility (Klatt, 1979). To summarise, it is clear that different prosodic factors can interact both with one another and with segmental factors to determine prosodic timing patterns, but the picture remains incomplete, especially as most of the studies reviewed in this section use highly controlled read sentences or nonsense words and only a few examine the concurrent effects of multiple influences in more naturalistic types of speech.

1.3. Relationships among phonetic parameters: The involvement of f0 and intensity in accentuation

Parameters that code accentual prominence are not all fully independent of one another. Fundamental frequency and intensity constitute a clear example of this phenomenon. Both play a role in coding prominence relationships in speech (Lieberman, 1960), but there is no general agreement in the literature as to which parameter might play a larger role, F0/pitch (e.g. Rietveld & Gussenhoven, 1985; Terken, 1991) or intensity/loudness (e.g. Kochanski, Grabe, Coleman, & Rosner, 2005). This disagreement might be, at least in part, explained by cross-linguistic differences in the weighting of acoustic cues to prominence (e.g. Barry, Andreeva, & Steiner, 2007). Moreover, F0 and intensity are known to co-vary during speech production (e.g. Gramming, Sundberg, Ternström, Leanderson, & Perkins, 1988; Hirano, Ohala, & Vennard, 1969; Tilsen, 2016), though the reasons for this frequently observed covariation are far from well understood. Physiological processes appear to contribute to the relationship between the two parameters, especially the changes in the subglottal pressure required during phonation and variation in loudness (Gramming et al., 1988; Hirano et al., 1969). A recent study has also demonstrated a great deal of inter-speaker variability in the strength of the correlation between F0 and intensity, suggesting that physiological factors alone do not suffice to fully account for the effect (Tilsen, 2016). The tight link between F0 and intensity is also measurable in perception (Melara & Marks, 1990), and the two acoustic parameters are assumed to be processed in an integral fashion (Grau & Kemler-Nelson, 1988). Most importantly, the covariation in F0 and intensity is present in the use of linguistic categories of prominence, i.e. pitch accents. Rosenberg & Hirschberg (2006) showed that the location of an intensity peak could help predict the type of pitch accent with 82% accuracy. Similar connections apply across other pairs of parameters, though they are less well explored. For example, rises take more time to be executed than falls, for physiological reasons (Xu & Sun, 2002) and delayed peaks produce higher pitch than early peaks (Gussenhoven, 2002; Rathcke, 2017).

The above considerations reinforce the idea that we cannot expect even well-established correlates of prosodic structure to be immune to complex variation. For example, if stressed and accented syllables are produced in a given dialect with a delayed f0 peak, intensity might pattern with f0, which would cause the accentually prominent syllable to have few physical prominence cues located on it. Such differences would, in the terms of (Nolan & Asu, 2009), reduce the dialect's *prominence gradient*: "in order to be prominent, to 'stick out', a syllable really needs to have more of the properties which lend prominence than do the syllables on either side, or at least one side" (Nolan & Asu, 2009). Alternatively, different cues might decouple, with (for example) increased duration and increased intensity being located on the stressed syllable, overriding the tendencies for duration and intensity to correlate with f0.

1.4. A cross-dialectal perspective on prominence and timing

Comparing dialects of a language offers a valuable window on the relationship between prosodic function and acoustic form. Dialects of a language share basic aspects of structure and function, but can differ in key segmental and suprasegmental constraints, which potentially affect the expression of the prosodic hierarchy. There is relatively little research in this area to date. White, Payne, and Mattys (2009) showed that Venetian and Sicilian varieties of Italian differ in prosodic timing patterns. They expected to find both greater prosodic lengthening and a generally sharper temporal prominence gradient in Sicilian, but observed only the former: the varietal differences appeared to be confined to phrase edges and/or accented positions. However, these two aspects of prosodic structure were not disentangled in their study. For English, Rathcke and Stuart-Smith (2016) have shown that segmental phonological structure affects prosodic timing patterns: under non-laboratory conditions, the working-class Glaswegian accent patterns with quantity languages, in that short vowels that undergo the Scottish Vowel Length Rule (see 1.6 for more detail) appear not to lengthen under accentual prominence (Rathcke & Stuart-Smith, 2016). A deeper exploration of these patterns is warranted, teasing apart effects of accentual prominence from those of finality, using a carefully controlled set of vowels.

The present study compared Standard Scottish English spoken with the urban accent of Glasgow (henceforth “Glasgow”) with Standard Southern British English (henceforth “SSBE”). These dialects differ in key aspects of both their segmental and intonational phonology, as described in detail in Sections 2.1.1 and 2.1.2. In outline, their vowel systems obey different constraints with respect to duration: Glasgow has the Scottish Vowel Length Rule (Agutter, 1988; Aitken, 1981; Scobbie, Hewlett, & Turk, 1999), whereas SSBE has a tense-lax distinction between a number of vowel pairs. In both cases, some kind of contrast between short and long vowels is involved, but the affected vowels and the phonological conditioning differ, and Glasgow vowels also tend to be shorter overall than SSBE ones. Moreover, the two dialects’ intonational systems are divergent: in Glasgow, the default intonational pattern on declarative sentences is a rise (or a delayed peak), whereas in SSBE, it is a fall. Thus, comparing these two dialects allows us to explore how both intonational and segmental constraints affect the realisation of accentual prominence, in interaction with phrase-finality, with implications for our theoretical understanding of these phenomena.

1.5. Research questions and hypotheses

We focused on trochaic words, which offer the scope to explore the domains of accentual lengthening in relation to that of phrase-final lengthening. With the above context in mind, we formulated the following research questions and hypotheses:

(1) What prominence gradients do trochaic words exhibit across two segmentally and intonationally distinct dialects of English? How do prominence gradients reflect lexical stress, and how are they modulated by accentual prominence and phrasal position?
Hypothesis (1): From perceptual impression and prior phonetic description, we predict less steep prominence gradients in Glasgow English across one or more of the parameters f0, intensity and duration.

(2) To what extent are accentual prominence and phrasal position expressed by similarly weighted and clustered phonetic properties in the two dialects, and to what extent are the two linguistic functions served by different acoustic means across the two dialects?
Hypothesis (2a): If prominence-related parameters are predominantly physiologically governed, we would expect to see similar relationships among prosodic parameters (pitch and duration, pitch and intensity) in the two dialects, even though these may translate into different surface patterns due to systemic prosodic and segmental differences.
Hypothesis (2b): If on the other hand the various properties involved in accentual prominence are not physiologically but primarily linguistically governed, we might see the parameters that contribute to accentuation and phrasing clustering differently across the dialects.

(3) To what extent are the acoustics of accentual prominence and phrasal position constrained by (dialect-specific) segmental phonology?
Hypothesis (3): If prominence involves “localized hyperarticulation” (de Jong, 1995) then a dialect’s systems of segmental phonological contrasts will constrain the expression of prominence. To maintain or hyperarticulate contrasts under prominence, contrastively long sounds will lengthen more than contrastively short sounds, with durational and spectral cues potentially trading off such that where lengthening cannot be used to express prominence, spectral correlates take over.

2. Method

2.1. Dialects

2.1.1. Segmental phonology: Key cross-dialect differences

Glasgow and SSBE have a set of striking differences in terms of their vowel systems. Glasgow has a smaller inventory, with nine monophthongs and three diphthongs compared to the 13 monophthongs and six diphthongs characteristic of SSBE (Abercrombie, 1979; Stuart-Smith, 1999, 2003, 2004; Wells, 1982: 364-5). The phonetic qualities of many vowels in the respective systems differ, e.g. Glasgow has monophthongal /e/ (FACE) and /o/ (GOAT) where SSBE has diphthongs /ɛɪ/, /əʊ/, respectively. The available evidence, though limited in terms of direct cross-dialect comparisons, suggests that /i u e o/ (among other vowels) are phonetically shorter in Glasgow than SSBE (Agutter, 1988; Ladd, 2005; McKenna, 1988; Wells, 1982).

Table 1
Typical patterns of vowel duration according to following phonological context in SSBE and Glasgow. Although the terms “short” and “long” suggest a binary distinction, vowel duration in fact increases successively from the top to the bottom row of the table, but the difference between successive “short” rows is small, whereas the difference between “short” and “long” rows is large (Scobbie et al., 1999).

Following context	Example	SSBE pattern	Glasgow pattern
Voiceless stops and fricatives	*seat*	short	short
Voiced stops, nasals and /l/	*seed*	long	short
Voiced fricatives and /r/	*seize*	long	long
Word or morpheme boundary	*see*	long	long

Glasgow and SSBE differ in the way vowel duration is conditioned by phonological context. SSBE operates according to the "Voicing Effect", i.e. vowels are longer before a voiced than a voiceless obstruent. Table 1 shows the schematic pattern for /i/: it is longer in *seed* than in *seat* and, as fricatives condition longer duration than stops, slightly longer still in *seize* (House & Fairbanks, 1953; Keating, 1985; Peterson & Lehiste, 1960). It is also longer when in an open than a closed syllable, and when the vowel directly precedes a word or morpheme boundary (e.g. *see;* Beckman & Edwards, 1990; Berkovits, 1994; Wightman et al., 1992). Table 1 shows a simplified view of the main patterns. In Glasgow, the Voicing Effect does not seem to apply, and vowel durations instead obey what is known as the *Scottish Vowel Length Rule* or *SVLR* (Agutter, 1988; Aitken, 1981; Scobbie et al., 1999). The SVLR applies to the high vowels /i/ and /u/ and to the diphthong /aɪ/ (Scobbie et al., 1999). When SVLR-vowels precede a voiced stop, nasal, or /l/, they are short: only slightly longer than they would be before voiceless consonants (McKenna, 1988; Scobbie et al., 1999). Only before voiced fricatives /v ä z ʒ/ and /r/ are the SVLR-vowels substantially longer than before voiceless consonants. Long duration is also conditioned by open syllables, word boundaries and morpheme boundaries, as in SSBE. An interesting consequence of SVLR in Glasgow is what Scobbie and Stuart-Smith (2008) term 'a quasi-phonemic contrast', i.e. a morphologically-conditioned distinction in a few word pairs. For example, in *brood—brewed,* the tautomorphemic /d/ of *brood* conditions short /u/, while the heteromorphemic /d/ of *brewed* conditions long /u/. SSBE lacks this distinction.

Glasgow and SSBE also differ in their use of the tense-lax contrast in vowels. This contrast is robust both qualitatively and quantitatively in SSBE, i.e. pairs such as /i/-/ɪ/, /ɛɪ/-/ɛ/, /ɑ/-/æ/, /ɔ/-/ɒ/, /u/-/ʊ/ differ in both quality and duration, with the first (tense or diphthongal) member of each pair being longer and more peripheral, and the second (lax) member of each pair being shorter and more central. In Glasgow, the tense-lax contrast lacks any clear application. The front vowels belong to phonemic categories that correspond to those of SSBE, i.e. /i/, /ɪ/, /e/ and /ɛ/. The back vowel pairs have undergone mergers in Glasgow (Abercrombie, 1979; Wells, 1982), so that /ɑ/ and /æ/ form a single category (henceforth /a/) as do /ɔ/ and /ɒ/ (henceforth /ɔ/) and /u/ and /ʊ/ (henceforth /u/). Thus some tense-lax pairs do not contrast due to mergers. For the remaining monophthongs, Ladd (2005) investigated whether a tense-lax opposition exists in Scottish English spoken in Edinburgh, and found no evidence of one in the vowels' durations themselves, nor did their behaviour with respect to f0 peak alignment or segmental compression give any indication of a covert tense-lax distinction. Height seemed to be the only factor possibly constraining duration. There has been no comparable investigation for Glasgow.

The systemic differences allow interesting predictions regarding the way prominence might be manifest durationally. Following de Jong (1995)'s framing of prominence as localised hyperarticulation, we can assume that short duration is critical to contrastiveness in SSBE lax vowels, and also in Glasgow /i/ in non-SVLR contexts, and possibly in some of Glasgow's other short vowels, e.g. /ɪ/. These vowels might therefore be limited in their ability to undergo prominence-related lengthening (Rathcke & Stuart-Smith, 2016). Vowels which are short before [-voice] consonants are likely to behave the same way, at least in SSBE. We might further expect those vowels that show limited accentual lengthening to show stronger spectral correlates of accentuation.

2.1.2. Prosodic phonology: Key cross-dialect differences

Glasgow and SSBE also differ in their intonational phonologies. Cruttenden (1997) proposed that in several 'Urban Northern British' cities (Glasgow along with Liverpool, Belfast, Newcastle, and Birmingham), the default realisation of a pitch accent – including a nuclear accent in a declarative sentence – is a rise (cf. Mayo, Aylett, & Ladd, 1997; Sullivan, 2011). These authors variously describe Glasgow phrase-final contours as having a low rise (Cruttenden, 1997), a 'rise-plateau' (i.e., a rising contour which stays level to the end of the phrase), or a 'rise-plateau-slump' (i.e., a rising contour where the peak is followed by a gradual fall; Mayo et al., 1997). Cruttenden (2007) observes these patterns in spontaneous speech in Glasgow, but not in read speech where falls (akin to SSBE patterns) are more common, a situation that he characterises as a type of "intonational diglossia". In contrast, SSBE obeys the typologically much more widespread pattern, where declarative nuclear accents are typically falls (Gussenhoven, 2004).

We expected that these intonational differences would affect the phonetic cueing of accentual prominence. Rises take longer to be executed (Xu & Sun, 2002), and delayed peaks produce higher pitch than early peaks (Gussenhoven, 2002; Rathcke, 2017). Therefore, we might expect to see an effect of pitch accent type on duration, with lengthening in the vicinity of a rise, compared to a fall. In this regard, it is interesting to consider how SSBE and Glasgow might differ in terms of their prominence gradients. Abercrombie (1979) impressionistically noted rhythmic differences between Standard Scottish English (SSE) and other varieties and proposed that in a trochee like *table,* SSE has a "short-long" rhythm where SSBE has an "equal-equal" rhythm. Unfortunately, the literature on rhythmic differences between the varieties is extremely scarce, and Abercrombie's proposal has not been tested quantitatively, except in a small corpus study by Rathcke and Smith (2011) which found strong syllables to be shorter, and weak syllables longer, in Glasgow than SSBE trochees. Measures of the prominence gradient applied across the syllables of trochees, for f0, intensity and duration, will help to clarify how intonation and timing are related.

2.2. Materials

To test the hypotheses outlined in Section 2, we manipulated two factors responsible for prosody-segment interactions: (1) vowel quality (tense or diphthongal vs lax vowels, for SSBE; low vs mid vs high for both dialects); (2) phonological context in the form of manner class of the word-medial consonant (voiceless stop vs. voiced fricative vs. /r/). The experimental materials are shown in Table 2, and comprise 48 trochaic words, ˈ(C(C))VC$_{\text{med}}$V, ending in *–y,* e.g. *cheesy, tricky, Katy, bevy*.

For SSBE, there were four pairs of vowels, each pair consisting of a tense vowel or diphthong, paired with a lax vowel: /i/-/ɪ/, /ɛɪ/-/ɛ/, /ɑ/-/æ/ and /ɔ/-/ɒ/. For Glasgow, the same words were used although the mid front pair is /e/-/ɛ/ (both monoph-

Table 2
Experimental words.

Medial consonant	Stressed vowel pair	Target words	
Voiced: • **fricative**	**/i/-/ɪ/**	cheesy breezy easy	chivvy privy Izzy
	SSBE /ɛɪ/-/ɛ/ **= Glasgow /e/-/ɛ/**	hazy daisy lazy	heavy bevy levy
• **rhotic**	**SSBE /ɔ/-/ɒ/** **= Glasgow /ɔ/**	Rory story Tory	lorry sorry Corrie
	SSBE /ɑ/-/æ/ **= Glasgow /a/**	Mairi tarry sari	marry carry Harry
Voiceless: • **stop**	**/i/-/ɪ/**	peaky creaky sleepy	picky tricky slippy
	SSBE /ɛɪ/-/ɛ/ **= Glasgow /e/-/ɛ/**	Haiti weighty Katy	Hetty yeti petty
	SSBE /ɔ/-/ɒ/ **= Glasgow /ɔ/**	naughty haughty gawky	knotty hockey dotty
	SSBE /ɑ/-/a/ **= Glasgow /a/**	harpy hearty party	happy Hattie patty

thongs), and each back pair is merged into a single category, /a/ and /ɔ/. C_{med} was either voiceless or voiced. Voiceless C_{med} were all voiceless stops, /p/, /t/ or /k/. Voiced C_{med} were voiced fricatives for the /i/-/ɪ/ and /ɛɪ/-/ɛ/ pairs, but /r/ was chosen as C_{med} for the back vowel pairs, as there are insufficient real words containing the back vowels followed by voiced fricatives.

As Table 2 shows, tense-lax minimal pairs were used (e.g. *easy/Izzy, peaky/picky*), wherever we could find suitable words from the ordinary English lexicon or phonologically plausible names. Where minimal pairs could not be found, words were paired with as close a phonological pairing as possible, with respect to the structure of the onset of the initial syllable, the place, manner and voicing of the consonant(s) of the initial syllable, and the place of articulation of C_{med}. Examples of non-minimal pairs are *cheesy/chivvy, breezy/privy, daisy/bevy, Rory/lorry*.

For the pair /ɑ/-/a/ followed by a voiceless C_{med}, the only available set of words with /ɑ/ had orthographic <ar>, e.g. *party, harpy*. These words are pronounced with /ɑ/ in the non-rhotic SSBE accent, but with /ar/ in rhotic Glasgow. We judged that the greatest comparability in analysing these words would be achieved by labelling a single sonorant interval (i.e., /ar/) for the Glasgow data.

The keywords were placed into sentences in three prosodic positions/conditions: ***n**uclear-accented, **n**on-final* (henceforth ***nn***); ***n**uclear-accented, utterance-**f**inal* (***nf***); ***p**ost-nuclear, utterance-**f**inal* (***pf***). Each sentence was designed to contain four trochaic Abercrombian feet (Abercrombie, 1967), in some cases with an anacrusis at the start. The keyword formed the last foot in ***nn*** and ***pf*** conditions, and the penultimate foot in the ***nn*** condition. Examples for the keyword *cheesy* are:

***nn**: That ˈsounds a ˈbit too ˈ**cheesy** ˈfor them*
nf**: I ˈreckoned ˈit was ˈkind of ˈ**cheesy
***pf**: I ˈdon't think ˈit's **re**ˈ**motely** ˈcheesy*

(stress mark indicates stressed syllable; bold font indicates nuclear-accented word). Crucially, this design allowed us to estimate the effect of accentual prominence, by comparing accented vs. non-accented words in final position (***nf*** vs. ***pf***). It also allowed us to estimate the effect of finality, by comparing nuclear-accented words in phrase-final vs. -medial positions (***nf*** vs. ***nn***).

To elicit the intended prosody in a natural manner, and especially to avoid the Glasgow speakers speaking in "read speech" mode which elicits SSBE-like prosody (Cruttenden, 2007) the sentences were embedded in scripted dialogues (examples in Appendix 1). Each dialogue consisted of four turns to be spoken by two speakers: ABAB. The sentence containing the keyword was always A's second turn, with A being the target speaker and B the confederate (see 4.2).

2.3. Participants

For each dialect, five participants (two male) were recorded as target speakers, all aged between 20 and 35 at the time of recording. The small number of participants for each dialect is offset to some extent by a relatively large number of tokens for each speaker and the careful control of the materials. No Glasgow participants had been resident in Southern England, and three of the SSBE participants had never resided in Glasgow. The other two SSBE participants, who were recorded in Glasgow, had been resident there for <3 years at the time of recording and had no detectable Scottish features of pronunciation or intonation, as judged by the first author, a native speaker of SSBE. All were educated speakers who read the dialogues fluently.

There were also two confederate dialogue partners. The confederate for the Glasgow participants was a 22-year-old female speaker of Glasgow English. The confederate for the SSBE participants was the first author, a native speaker of SSBE and 35 years old at the time of the recordings.

2.4. Procedure

The order of the dialogues was randomised within each prosodic condition. The dialogue partner (confederate, role B) and participant (role A) each had a script. The dialogue partner's copy was annotated with the intended prosody for the participant's key turn. If the participant produced the utterance disfluently, or with an error affecting the prosody or segments of the keyword, the dialogue partner invited them to repeat the dialogue, which in most cases resolved the error or disfluency.

SSBE participants F1, F2 and M1 were recorded in the sound treated studio of the Phonetics Laboratory at the University of Cambridge. SSBE participants F3 and M1, and all Glasgow participants, were recorded in the sound-treated studio of the Glasgow University Phonetics Laboratory. A Sennheiser MKH40 P48 condenser microphone placed about 30 cm from the participant's mouth was used. The participant and their dialogue partner recorded the entire set of dialogues three times. The recordings took between 2 and 3 hours, including breaks.

2.5. Data preparation

Tokens where the keyword had a segmental mispronunciation, was affected by disfluency/hesitation, or was produced with the wrong prosody (i.e. with accentuation where none was intended, or vice versa) were excluded from the dataset. Tokens were also excluded if the foot structure on or after the keyword was not produced as intended (e.g. if the participant said *That ˈsounds a ˈbit too ˈ**cheesy** for ˈthem,* with *cheesy* in a trisyllabic foot, instead of the intended *That ˈsounds a ˈbit too ˈ**cheesy** ˈfor them).* Deviations from the intended trochaic rhythm that occurred earlier in the prosodic phrase were ignored (e.g. *ˈThat sounds a ˈbit too ˈ**cheesy** ˈfor them*). Out of a possible 4320 tokens (48 words × 3 prosodic positions × 3 repetitions × 10 speakers), 4015 were analysable, i.e. 9.6% of tokens were discarded.

Labelling was carried out by a trained phonetician, and approximately 25% of the data were checked by the first author. The keyword was segmented into its constituents as follows (example in Fig. 1): one or more initial consonants (*ini*); stressed vowel (*svo*); medial consonant (*med*); and unstressed vowel (*uvo*). Segmentation was carried out according to standard criteria (e.g. Turk, Nakai, & Sugahara, 2006). For stops, closure duration and VOT were labelled separately. For fricatives, the start and end of aperiodic noise were used. For liquids, the main F2 transitions and spectral discontinuity were used; the start and end of the consonant interval were labelled at the beginning and end of the F2 transition, respectively. For utterance-final vowels, the end of the vowel was marked at the end of voicing. For vowels, we additionally labelled pre-aspirated, glottalised and breathy intervals. As noted in Section 2.2, the sequence /ar/ before a stop in Glasgow was labelled as a single sonorant interval for comparability with non-rhotic SSBE's /ɑ/ (e.g. *hearty:* /hɑti/ in SSBE, /harti/ in Glasgow). For obstruent consonants, we additionally labelled voiced/devoiced intervals that contrasted with the obstruent's phonological specification; and lenition (fricative realisations of stops, and approximant realisations of fricatives). For /r/, we labelled tap and trill realisations. These are not analysed further here.

2.6. Acoustic measurements

The following set of acoustic parameters was measured: durations of stressed and unstressed vowels, F1/F2, F0 and amplitude at the temporal midpoint of the stressed and the unstressed vowel. All measures were taken in Praat.

For each word, we then derived further measures to normalise the data and/or to capture relationships between the stressed and the unstressed vowel (or its substitute as appropriate). These measures include:

(1) Vowel quality of stressed and unstressed vowels, expressed as the Euclidean distance in Hz between the vowel's position in F1 ~ F2 space and a speaker-specific centroid, defined as the average F1 ~ F2 of the most peripheral vowel qualities of our corpus, /i/, /ɔ/ and /ɑ/;
(2) F0 magnitude in st following the equation 12*log2(F0stressed/F0unstressed) which represents falls and rises from the stressed to the unstressed syllable as positive and negative values, respectively;
(3) Intensity ratio: the difference between the intensity values of the stressed and unstressed syllable in dB, derived through subtraction;

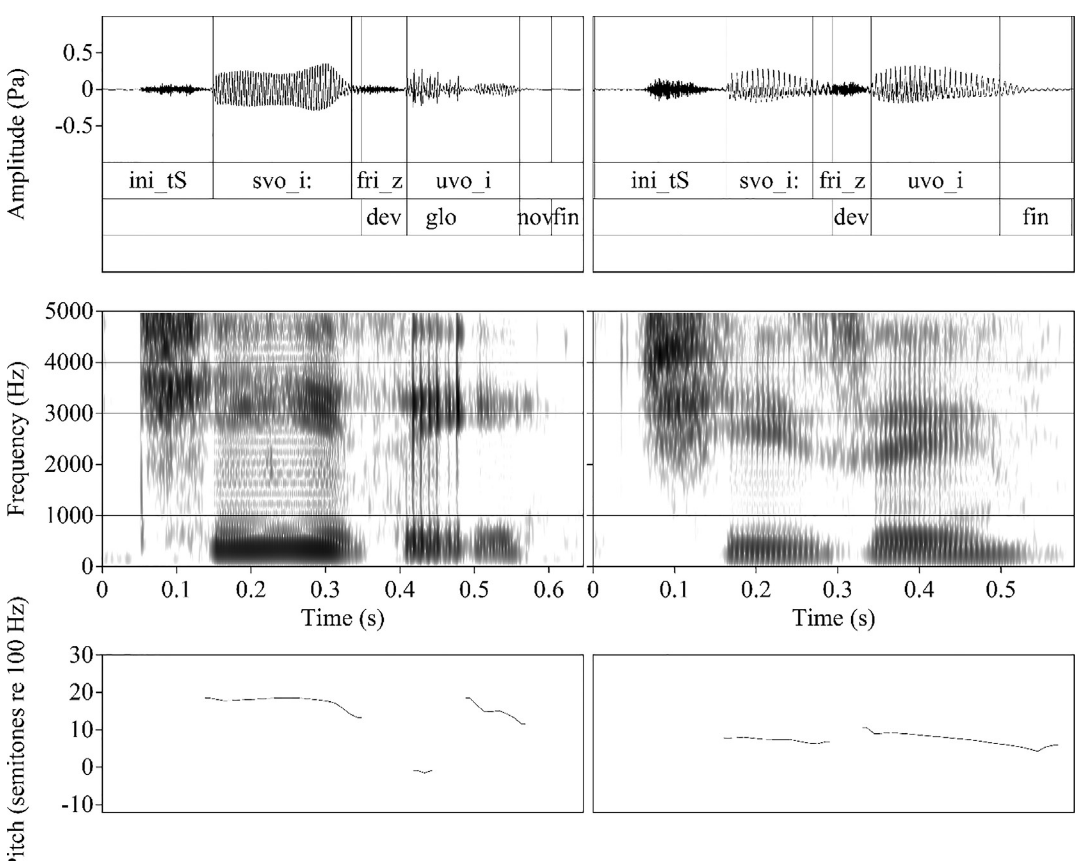

Fig. 1. Labelled tokens of *cheesy* in *nuclear final* position, from a female SSBE speaker (left) and a female Glasgow speaker (right).

(4) Durational ratio, calculated by dividing the duration of the stressed vowel by that of the following unstressed vowel in the same word. Values close to 1.0 indicate a similar duration across the two vowels, and a smaller prominence gradient.

Across all of the ratio measures in (2)–(4), a larger ratio value indicates a larger acoustic prominence gradient.

2.7. Data analyses and modelling

Analyses were conducted in RStudio (version 1.0.136). Where linear mixed effects models were used, they were fit using the lme4 library, with REML t-tests using Satterthwaite approximations to degrees of freedom. The model fitting procedure was as follows. First, for each predictor, a saturated model with all relevant main effects and interactions was fitted. Next, the step() function was used to remove non-significant predictors. Planned comparisons were obtained using the output of step() where available. For three-way or higher-order interactions, step() does not output planned comparisons, so we obtained these by re-levelling the initial model. R code for the final models is in Appendix 2.

In every model, we tested the factor *prosody* (with three levels, ***nn***, ***nf*** and ***pf***, of which only the ***nn/nf*** and ***nf/pf*** comparisons were analysed, resulting in an adjusted alpha-level of 0.025), the factor *dialect* (levels: SSBE and Glasgow), and their interaction. Additionally, variables relating to stressed vowel quality were defined for each dialect separately: *tenseness* for SSBE (two levels: *tense, lax*), and *vowel identity* for Glasgow (six levels: /i/ /ɪ/ /e/ /ɛ/ /a/ /ɔ/), and the variable *medial consonant voicing* was used in some analyses. A list of covariates was defined for each model as appropriate (see Section 4). *Speaker* and *word* were treated as random effects (intercepts only, since including slopes led to failures to converge).

To answer research question (1) *(What prominence gradients do trochaic words exhibit across two segmentally and intonationally distinct dialects of English? How do prominence gradients reflect lexical stress, and how are they modulated by accentual prominence and phrasal position?)* we modelled three dependent variables, f0 magnitude, intensity ratio, and duration ratio, as a function of *prosody, dialect* and their interaction, plus covariates in each case as outlined in Section 4.

Research question (2) (*To what extent are accentual prominence and phrasal position expressed by similarly weighted and clustered phonetic properties in the two dialects, and to what extent are the two linguistic functions served by different acoustic means across the two dialects?)* was addressed using a dimension reduction technique. Principal Components Analysis (PCA) allows complex multidimensional datasets to be reduced to their main axes of variation, and in doing so, reveals how the dimensions of the original dataset co-vary. By exploring the “loadings” of the original variables on the principal components, it is possible to discover how properties cluster together to determine the structure of the data. We ran PCAs using dependent variables as follows: (1) *stressed vowel duration* (in ms), (2) *unstressed vowel duration* (in ms), (3) *initial consonant duration* (in ms), (4) *medial consonant duration* (in ms), (5) *stressed vowel peripherality* (Euclidean distance as defined in 3.5 above), (6) *unstressed vowel peripherality*, (7) *stressed vowel f0* (in Hz), (8) *unstressed vowel f0* (in Hz), (9) *stressed vowel pitch* (in st relative to the speaker’s lowest f0 as baseline), (10) *unstressed vowel pitch* (in st), (11) *f0 magnitude* (in st), (12) *duration of f0 change* (in ms), (13) *speed of f0 change* (in st/sec), (14) *stressed vowel intensity*, (15) *unstressed vowel intensity*, (16) *intensity ratio*, (17) *duration ratio,* and (18) *number of segments*. All predictors were scaled and centred. Separate PCAs were carried out for each dialect. This was because the dialects could, in theory, differ in two ways: (1) the principal components might be associated with different combinations of variables in each dialect, and/or (2) the main predictors in our design might affect the principal components in different ways. Separate PCAs offer more insight into the first of these questions. Accordingly, we established the principal components for each dialect. We then ran linear mixed effects models on the values of the first four principal components as a function of *prosody*, to reveal the effect of this design variable on each of them.

To answer research question (3) (*To what extent are the acoustics of accentual prominence and phrase-finality constrained by (dialect-specific) segmental phonology?)* we modelled stressed vowel duration, stressed vowel quality, and unstressed vowel duration as a function of *prosody, dialect,* and variables relating to the dialect’s vowel system. These last were defined for each dialect separately: *tenseness* for SSBE (two levels: *tense, lax*), and *vowel identity* for Glasgow (six levels: /i/ /ɪ/ /e/ /ɛ/ /a/ /ɔ/). The variable *medial consonant voicing* was used in some analyses, along with covariates as described in Section 4. For the analysis of /i/ in Glasgow, *medial consonant voicing* is more properly thought of as *SVLR-conditioning context,* since voiced consonants other than fricatives and /r/ do not trigger SVLR. The word’s *number of segments* and the *stressed vowel pitch* were included as covariates where they contributed to the model, but their effects are not reported below.

3. Results

3.1. Prominence gradients

This Section addresses the first research question, whether or not two segmentally and intonationally distinct dialects of English would exhibit different acoustic prominence gradients, and how those gradients might reflect lexical stress, accentual prominence and phrasal position. We turn to each acoustic parameter separately.

3.1.1. F0

The pitch pattern across the trochaic words was dramatically different across the dialects. As expected, the predominant pitch pattern was a fall from the stressed to the following unstressed syllable in SSBE (64%), and a rise in Glasgow (61%). 85% of these falls in SSBE and 88% of the rises in Glasgow were accentual. Post-nuclear final positions mostly showed either a flat trajectory or the dialect-untypical pitch pattern (i.e. in Glasgow, a fall, high in the speaker’s range; and in SSBE a rise, low in the speaker’s range). A χ^2-test showed that the distribution of rising, falling and flat pitch patterns heavily depended on the variety ($\chi^2(2) = 589.86$, $p < 0.001$).

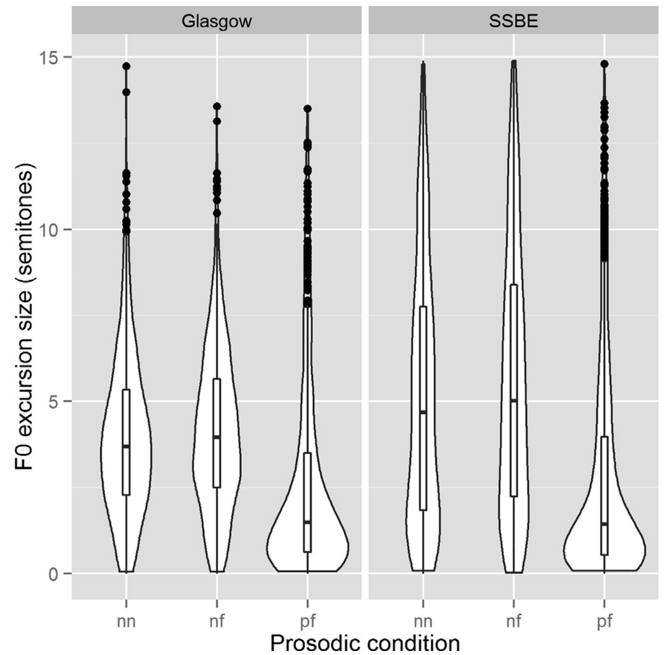

Fig. 2. F0 excursion size by dialect and prosodic condition. Violins show the data distribution as mirrored density plots. The superimposed boxplots show the median (thick line), first and third quartiles (box edges), and the smallest and largest values that are no further than 1.5× interquartile range from the quartiles (whisker edges), with outliers plotted as dots.

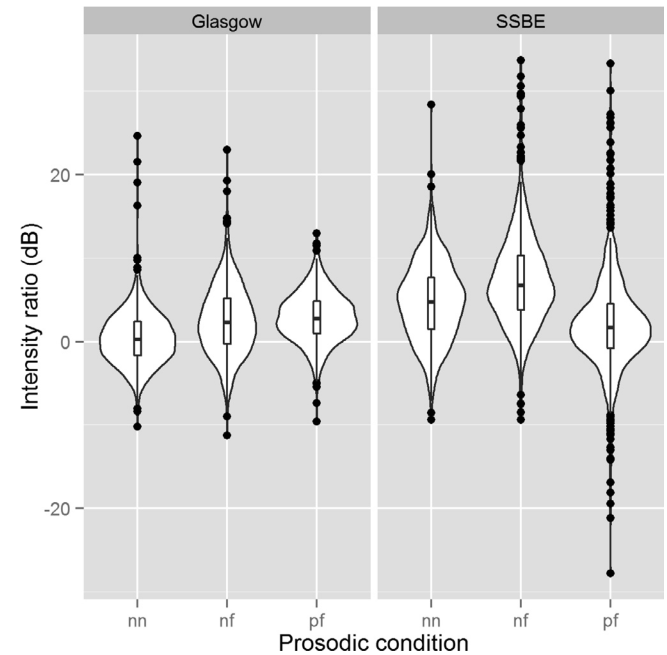

Fig. 3. Intensity ratio by dialect and prosodic condition.

Fig. 2 shows that the absolute excursion size of F0 changes between stressed and unstressed syllables was on average 1 st larger in SSBE (mean: 4.6 st) than Glasgow (mean: 3.6 st). Statistical modelling revealed a significant interaction of *dialect* and *prosody* ($F = 24.6$, $p < 0.001$). However, planned comparisons did not show significant differences between SSBE and Glasgow for any of the prosodic positions. Where the dialects did differ was that SSBE had slightly but significantly larger pitch excursions in *nuclear final* than *nuclear non-final* position (mean difference 0.5 st; $t(3872) = 3.18$, $p = 0.002$), whereas Glasgow's pitch excursions did not differ between these two positions (mean difference 0.1 st; $t(3871) = 0.93$, $p > 0.05$). This might reflect a tendency towards intonational compression in SSBE vs. truncation in Glasgow. As expected, both dialects, especially SSBE, had much larger excursions in *nuclear final* than *post-nuclear final* position (SSBE: 2.9 st, $t(3873) = 19.52$, $p < 0.0001$; Glasgow: 1.5 st, $t(3873) = 9.59$, $p < 0.0001$).

In sum, our pitch analyses demonstrate that accentual prominence in nuclear positions of declarative sentences is cued predominantly by falling pitch in SSBE, and by rising pitch in Glasgow.

3.1.2. Intensity

Intensity ratios are shown in Fig. 3: the more that stressed syllables exceed unstressed syllables in intensity, the higher the intensity ratio. Statistical results show that the intensity ratio is strongly dependent on dialect, as reflected in significant effects of *dialect, prosody* and their interaction, with *stressed vowel pitch* level and the word's *number of segments* significant as covariates.

In SSBE, the intensity ratio is generally positive, i.e. stressed vowels have higher intensity than unstressed vowels. Accentual prominence considerably boosts the intensity ratio (*nuclear final* > *post-nuclear final,* $t(3890) = 17.41$, $p < 0.0001$), and phrase-finality also increases it (*nuclear final* > *nuclear non-final,* $t(3871) = 12.75$, $p < 0.0001$). Note the large spread of positive ratios in the *nuclear final* condition: the extreme values mainly reflect cases where the final syllable was fully devoiced. In *postnuclear final* position there are both positive and negative extreme values, which reflect that either vowel in the word could be realised with very low intensity due to devoicing or creak.

In Glasgow, on the other hand, the vowels' intensities are fairly balanced (close to 0) in *nuclear non-final* position; only in final positions does the stressed vowel have higher intensity than the unstressed vowel. Accentual prominence does not increase the intensity ratio: there is no difference between the *nuclear final* and *post-nuclear final* positions ($t(3873) = 0.05$, n.s.). But the intensity ratio is greater in *nuclear final* than *nuclear non-final* position ($t(3868) = 9.27$, $p < 0.0001$).

Cross-dialectal comparisons showed that Glasgow has a significantly smaller intensity ratio than SSBE under accentual prominence, i.e. in *nuclear final* and *nuclear non-final* positions ($t(3) = 3.63$, $p < 0.01$; $t(3) = 2.90$, $p < 0.05$), but the dialect difference is non-significant in *post-nuclear final* position. In both dialects, the intensity ratio increases in proportion to the pitch level of the stressed vowel ($F = 65.95$, $p < 0.001$). That is, the higher the F0 measured in the centre of the stressed vowel, the higher the intensity ratio.

3.1.3. Duration

Dialect-specific distributions of the durational ratios are shown in Fig. 4. Statistical modelling again revealed significant effects of *dialect, prosody* and their interaction, along with *stressed vowel pitch* and the word's *number of segments* as covariates. Fig. 4 shows that the primary influence on durational ratios was phrase-finality, and dialect differences played a smaller role. Words in *nuclear non-final* positions show a

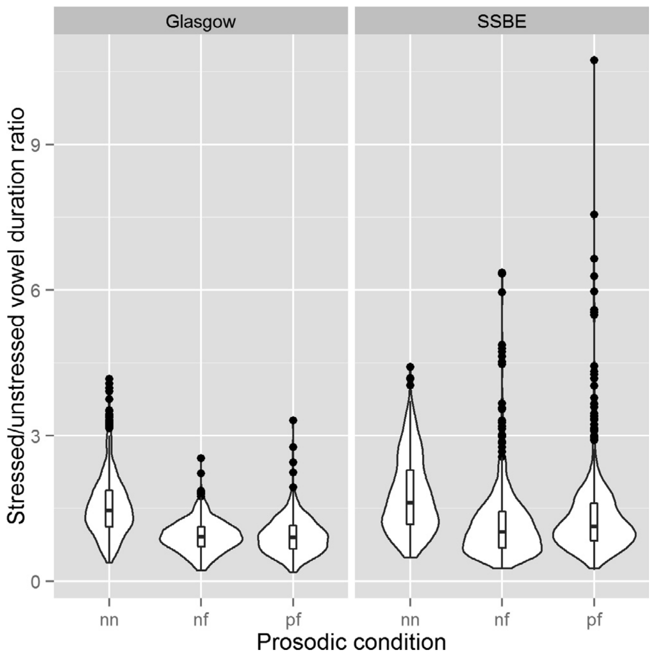

Fig. 4. Vowel duration ratio by dialect and prosodic condition.

durational ratio well above 1, meaning that their stressed vowel is longer than their unstressed vowel. In contrast, when words are phrase-final the two vowels' durations are more balanced (ratios close to 1). In both dialects, the ratios were greater in the *nuclear non-final* than in the *nuclear final* position (SSBE: $t(3867) = 19.45$, $p < 0.0001$; Glasgow: $t(3867) = 20.87$, $p < 0.0001$).

Accentual prominence had rather little effect on the measured durational ratios. Notably, accentual prominence did not enhance the contrast between stressed and unstressed vowel durations. On the contrary, in SSBE, the stressed/unstressed vowel duration ratio was slightly smaller in *nuclear final* than *post-nuclear final* words ($t(3867) = 4.89$, $p < 0.0001$) while in Glasgow, the ratio was unaffected by accentuation ($t(3867) = 0.2$, n.s.).

Glasgow had a significantly smaller stressed/unstressed vowel duration ratio than SSBE in *post-nuclear final* position ($t(3) = 2.66$, $p < 0.05$), while dialect differences were non-significant in other positions. The ratio decreases as the number of segments in the word increases, consistent with intra-syllabic compression of vowels in more complex syllables (cf. Katz, 2012, Munhall, Fowler, Hawkins, & Saltzman, 1992).

3.2. Principal components analysis

With the PCAs, we took a data-driven approach to the patterning of acoustic parameters (duration, pitch, intensity and vowel quality; specific parameters are listed in Section 3.5). We explored the variable loadings on the first four principal components per dialect, and conducted regression analyses to test how each component was affected by *prosody.* This analysis aimed to answer the second research question of the study, to what extent prominence and phrasal position are expressed by similarly weighted and clustered phonetic properties in the two dialects.

The basic structure of component loadings was similar across the two dialects: *magnitude of pitch change* was associated with the first principal component, *unstressed vowel duration* dominated the second, *stressed vowel f0* the third, and the *duration of stressed vowels* predominantly loaded on the fourth. (Further PCs, not explored here, might relate to consonantal parameters.) Thus a reasonable degree of similarity in the data structure exists regardless of whether a dialect's typical nuclear accent is a fall or a rise.

Nevertheless, there were substantial cross-dialect differences too. In particular, SSBE had a tighter relationship than Glasgow between the acoustic parameters under investigation and the prosodic dimensions above. For example, in the loadings on PC1 for SSBE, falling pitch and falling intensity correlated strongly, while in the regression analyses, PC1 was strongly increased by accentual prominence (*nuclear final* vs. *post-nuclear final* position, $t(3816) = 23.78$, $p < 0.0001$) and was unaffected by finality. That is, PC1 for SSBE clearly reflects closely clustered pitch and intensity changes associated with accentuation. The loadings for PC1 for Glasgow showed a much looser alignment of pitch and intensity, and a stronger influence of the unstressed vowel duration. The values of PC1 were significantly decreased by accentuation, (*nuclear final* vs. *post-nuclear final,* $t(3749) = 10.2$, $p < 0.0001$) reflecting larger f0 rises in accented than non-accented positions (cf. Section 4.1.1) and increased by finality (*nuclear final* vs. *nuclear non-final,* $t(3744) = 23.6$, $p < 0.0001$).

For other principal components too, the parameters clustered differently across the dialects. The values of PC2 for both dialects strongly reflected finality and secondarily accentual prominence; but the loadings on this component were different for the two dialects: for Glasgow PC2 reflected a conjunction of properties of the unstressed vowel—long duration, high pitch, and high intensity, whereas for SSBE PC2 was dominated by the unstressed vowel's duration only. Conversely, PC3 was dominated by the f0 of the stressed vowel for both dialects, but for SSBE this was strongly correlated with the stressed vowel's intensity and duration, whereas for Glasgow it was not. Finally, the fourth principal component was dominated by stressed vowel duration, but this was more strongly correlated with stressed vowel peripherality in SSBE than in Glasgow.

3.3. Prosody-segment interactions

The final section seeks to provide evidence relating to the third research question of this study, how the segmental constraints within a system affect the realisation of prominence and finality. To do this, we ran regression analyses on individual acoustic parameters: duration and peripherality of the stressed vowels, and duration of unstressed vowels.

3.3.1. Timing of stressed vowels

We ran separate models for SSBE and Glasgow, because their stressed vowel systems obey different constraints, and two separate models for Glasgow (/i/ vs. all other vowels) for similar reasons. In other respects, the initial models were the same for both dialects, and included prosodic condition and medial consonant voicing and their interaction, together with number of segments within a syllable and the stressed vowel pitch as covariates. Intrasyllabic compression and vowel lengthening in higher pitch were observed in both dialects, though Glasgow data showed these effects exclusively in the SVLR-vowel /i/.

Crucially, the best-fit model for SSBE showed that the durational effects of prosodic condition interacted with segmental constraints, reflected in the three-way interaction of *tenseness* with *prosody* and *medial consonant voicing*. We observed accentual lengthening, but its magnitude was modulated by segmental constraints. More specifically, Fig. 5 shows that tense vowels followed by [+voice] segments lengthened significantly under accentuation (by 19 ms; $t = 9.19$, $p < 0.0001$), and tense vowels followed by [−voice] segments showed weak lengthening (4 ms, $t(1933) = 2.04$, $p < 0.05$). Lax vowels followed by [+voice] segments showed merely a trend towards accentual lengthening (4 ms; $t(1935) = 1.87$, $p = 0.06$) and lax vowels followed by [-voice] segments did not lengthen significantly under accentuation (1 ms; $t(1932) = 0.7$; n.s.). In summary, factors that keep vowels short—lax phonological status and an unvoiced following consonant—seem to limit the scope for accentual lengthening, in a cumulative manner. In contrast, phrase-final lengthening applied to all stressed vowels across the board, with only slight differences in the magnitudes of the lengthening effect (5–10 ms, all $t > 2.99$, all $p < 0.005$).

For Glasgow, stressed vowel durations are shown in Fig. 6. We analysed /i/ (which is subject to the SVLR) separately from the non-SVLR vowels. In the best-fit model, /i/ showed no significant accentual lengthening (3 ms; $t(229) = 1.21$, n.s.), but did show phrase-final lengthening (12 ms; $t(230) = 5.59$, $p < 0.0001$). Independently of prosodic condition, /i/ also showed a substantial effect of SVLR: it was 42 ms longer when followed by a voiced fricative than by a voiceless stop ($t(3) = 14.88$, $p < 0.001$).

As far as the Glasgow non-SVLR vowels are concerned, the best-fit model contained *prosody*, *medial consonant voicing, vowel identity*, and a three-way interaction among these

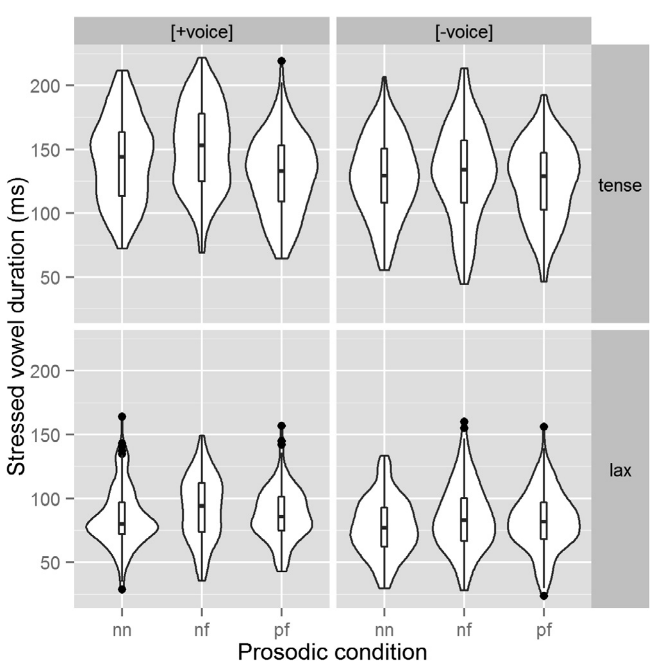

Fig. 5. Stressed vowel duration for SSBE, by prosodic condition, tenseness and medial consonant voicing.

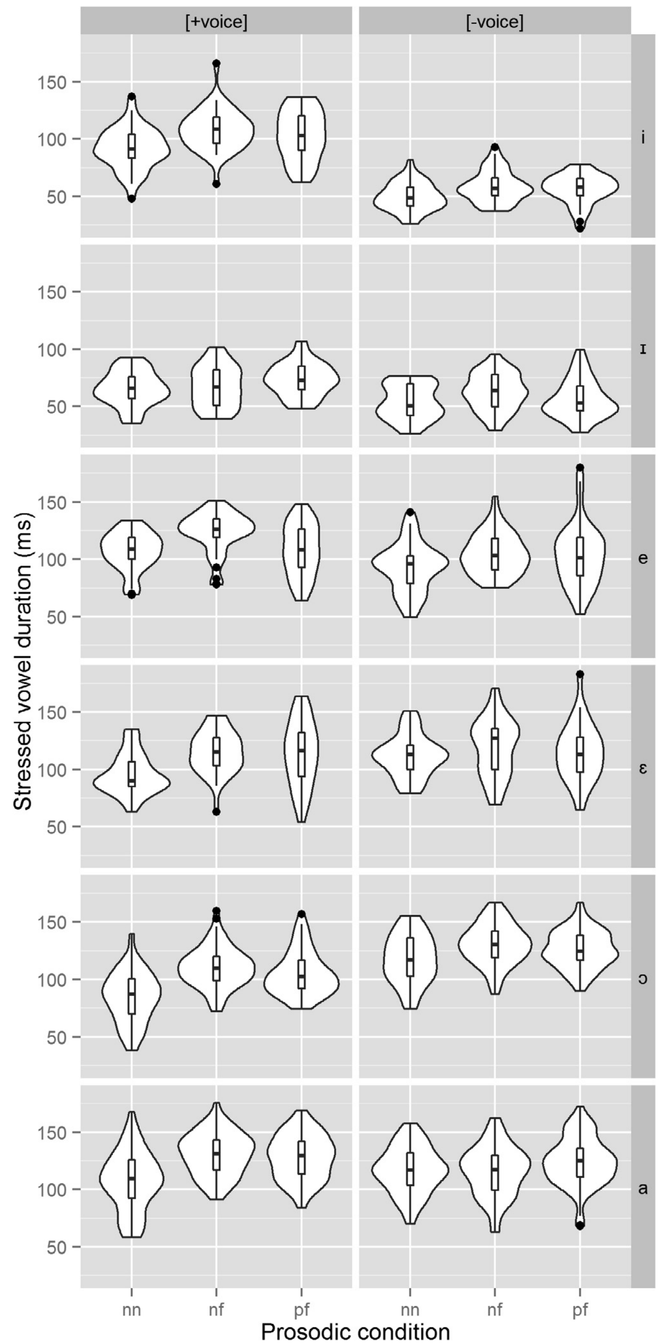

Fig. 6. Stressed vowel duration for Glasgow, by prosodic condition, vowel category and following segment voicing.

variables. Crucially, effects of *prosody* were shaped by the segmental constraints (Fig. 6). Accentual lengthening was only significant for the following two vowels when followed by a voiced consonant: mid-vowel /e/ (14 ms accentual lengthening, $t(1487) = 4.03$, $p < 0.0001$), and back-vowel /ɔ/ (5 ms, $t(1492) = 1.97$, $p < 0.05$); and for low-vowel /a/ when followed by a voiceless consonant (9 ms accentual lengthening, $t(1492) = 3.35$, $p < 0.001$). No other vowel showed accentual lengthening (/ɪ/ followed by a voiced consonant showed a trend, $p = 0.07$; all other estimates were 5 ms or less, all ts < 1.2, all ps n.s.). Thus, while the details were more complex than those found in the SSBE data, accentual lengthening responded sensitively to segmental constraints in Glasgow as well.

Similarly to SSBE, phrase-final lengthening was more widespread and less sensitive to segmental constraints. Seven out of ten combinations of vowel quality and following segment voicing that were tested showed significant lengthening of the stressed vowel in phrase-final positions (with magnitudes ranging from 9 ms to 26 ms, all ts > 2.6, all *p*s < 0.01). The exceptions were /ɛ/ before [−voice] consonants, /ɪ/ before [+voice] consonants and /a/ before [−voice] consonants; we could not detect a pattern in this particular vowel set.

3.3.2. Quality of stressed vowels

Two separate models were run for SSBE and Glasgow on the Euclidean distances from a speaker-defined centroid (see 3.5 for more detail). For SSBE, the best-fit model included only main effects of *prosody* and *vowel identity*. Accentual prominence and phrase-finality both affected peripherality: vowels in the *nuclear final* position were further from the speaker's centroid than those in the *postnuclear final* condition (by 40 Hz, $t(3812) = 4.24$, $p < 0.0001$) and also than those in the *nuclear non-final* position (by 24 Hz, $t(1811) = 2.53$, $p < 0.025$). Vowel category unsurprisingly influenced peripherality, in line with the qualities of the vowels in the dataset (results not reported further).

For Glasgow, the final model also included the interaction of *prosody* and *vowel*. The effect of accentual prominence on peripherality was rather large overall (116 Hz), but it depended on vowel type: /i/, /ɛ/, /a/ and /ɔ/ were all more peripheral in accented (*nuclear final*) than in non-accented (*postnuclear final*) position, but /e/ and /ɪ/ showed no spectral correlate of accentuation. This is particularly interesting for /e/, which was one of the few vowels that did show a durational correlate of accentuation. Phrase-finality did not increase peripherality in Glasgow.

3.3.3. Timing of unstressed vowels

The best-fit model included three two-way interactions (*dialect* and *prosody*: $F = 72.61$, $p < 0.001$; *dialect* and *medial consonant voicing*: $F = 14.97$, $p < 0.001$; *prosody* and *medial consonant voicing*: $F = 23.36$, $p < 0.001$), plus *pitch height in the stressed vowel* as a covariate ($F = 37.65$, $p < 0.001$).

Accentual prominence was expressed on the unstressed vowel differently in the dialects, as shown by the significant interaction of *dialect* and *prosody*. In SSBE, comparing across *nuclear final* and *post-nuclear final* words, the unstressed vowel showed small but significant lengthening when the word was accented (15 ms; $t(3880) = 11.12$, $p < 0.0001$). This indicates that in SSBE, accentual lengthening spreads beyond the stressed syllable to the unstressed one. In contrast, Glasgow unstressed vowels showed no such lengthening (*nuclear final* vs. *post-nuclear final:* $t(3868) = 0.63$, n.s.). Both dialects showed large, significant phrase-final lengthening (SSBE: 47 ms; $t(3868) = 37.39$, $p < 0.0001$; Glasgow: 55 ms; $t(3868) = 44.87$, $p < 0.0001$). Additionally, Glasgow unstressed vowels were longer than SSBE unstressed vowels, but only in *post-nuclear final* position (29 ms; $t(8) = 3.04$, $p < 0.025$).

The significant interaction of *dialect* with *medial consonant voicing* reflects that the voicing status of the previous consonant had no effect on the unstressed vowel duration for SSBE, but for Glasgow, vowels were slightly longer after [+voice] than [−voice] consonants (6 ms; $t(54) = 2.24$, $p < 0.05$).

The significant interaction of *voicing* with *prosody* reflects that accentual lengthening was stronger for unstressed vowels preceded by [+voice] than [−voice] consonants (11 ms, $t(3874) = 8.22$, $p < 0.0001$, vs 4 ms; $t(3868) = 3.2$, $p = 0.001$). This suggests that voiceless consonants may block the spread of accentual lengthening. In contrast, final lengthening was significant for vowels preceded by both [+voice] and [-voice] consonants. Finally, the duration of the unstressed vowel increased by a small but significant amount with an increase in the pitch of the stressed vowel (0.7 ms per semitone increase; $t(3901) = 6.14$, $p < 0.0001$).

3.4. Summary

Table 3 summarises the key findings for all three research questions, focusing on cross-dialectal differences.

4. Discussion and conclusions

This study investigated what a comparison of two intonationally and segmentally distinct varieties of English could reveal about the phonetics of accentual prominence. We sought to establish how the specific phonologies of SSBE and Glasgow English affect the domain of prominence-related cues and/or give rise to trading relationships among cues. Building on de Jong (1995: 502)'s argument that "stress [accentuation, in our terms] can act as a diagnostic for determining the content of the linguistic code of a particular language", we proposed that comparing varieties whose linguistic codes are known to differ might help us to understand the phonetic correlates of accentuation. The study revealed robust cross-dialect differences across a number of measures, though it is small-scale in terms of the number of participants and would ideally be replicated with a larger group.

First, we sought to establish the prominence gradients that trochaic words exhibit in the two varieties, asking how these gradients express lexical stress, and how they are modulated by accentual prominence and phrasal position. Hypothesis 1 predicted shallower gradients for Glasgow than SSBE: strong differences were confirmed in some prosodic positions, and SSBE gradients were never shallower than Glasgow gradients. A major difference between the dialects was that accentuation did not alter either the intensity or duration gradient for Glasgow. In contrast, accentual prominence sharpened the intensity gradient, but flattened the duration gradient for SSBE. The former pattern is expected, the latter unexpected, though perhaps not entirely unsurprising in light of Abercrombie (1979)'s comments on an "equal-equal" rhythm in SSBE trochees. It may reflect a degree of rightwards spreading of accentual lengthening from the stressed to the unstressed vowel, which we also observed in the regression analyses on unstressed vowel duration (4.3.2); or it could have a structural reason relating to ambisyllabicity, which we return to below. Finality has more consistent effects across the dialects: it flattens the duration gradient for both, as a consequence of phrase-final lengthening. Overall, the differences observed involve less sharp prominence gradients for Glasgow, and

Table 3
Summary of results, focusing on key cross-dialectal differences.

Research question	Prediction	Key results
(1) What prominence gradients do trochaic words exhibit across two segmentally and intonationally distinct dialects of English? How are prominence gradients modulated by accentual prominence and phrasal position?	Less steep prominence gradients in Glasgow than SSBE across one or more of f0, intensity and duration	*Broadly supported:* • F0: Similar prominence gradient for SSBE and Glasgow, but Glasgow has mainly accentual rises, SSBE falls • Intensity: Steeper prominence gradient for SSBE than Glasgow • Duration: Similar prominence gradients for SSBE and Glasgow, except in postnuclear final position, where SSBE is steeper than Glasgow • Accentuation increases prominence gradient for f0 (both dialects) and intensity (SSBE only) but decreases prominence gradient for duration (SSBE only) • Phrase-finality increases prominence gradient for f0 (SSBE only) and intensity (both dialects), but decreases prominence gradient for duration (both dialects)
(2) To what extent are accentual prominence and phrasal position expressed by similarly weighted and clustered phonetic properties in the two dialects?	Dialects should show similar relationships among prosodic parameters if these are predominantly physiologically governed; or different relationships if they are predominantly linguistically governed	*Linguistic view supported:* There are cross-dialectal differences in the weighting and clustering of acoustic parameters: • SSBE: Falling pitch and falling intensity are strongly correlated and reflect accentuation (PC1). Stressed vowels' pitch, intensity, duration and peripherality tend to correlate (PC3, PC4). Correlations among parameters for unstressed vowels are weaker (PC2) • Glasgow: Rising pitch reflects accentuation, correlating less with intensity and more with duration (PC1). Unstressed vowels' duration, pitch and intensity tend to correlate (PC2, reflecting finality). Correlations among parameters for stressed vowels are weaker (PC3, PC4). These patterns suggest a role for linguistic control factors
(3) To what extent are the acoustics of accentual prominence and phrasal position constrained by (dialect-specific) segmental phonology?	Within a system, contrastively long sounds will lengthen more than contrastively short sounds. Trade-offs between durational and spectral correlates of accentuation may occur.	*Supported:* Substantial cross-dialectal differences in the correlates of accentuation systematically reflect segmental constraints: • For SSBE, accentual lengthening occurs for tense vowels and vowels followed by [+voice] segments. Accentual lengthening spreads to the following unstressed vowel. Accented vowels are more peripheral than unaccented • For Glasgow, accentual lengthening is sporadic, occurs only on a few vowels, and does not spread to the following unstressed vowel. Accented vowels are more peripheral than unaccented; there is some evidence of trade-offs between peripherality and duration. Phrase-final lengthening appears much less sensitive to dialect-specific segmental constraints

the clearest contributing factors are that Glasgow does not mark accentual prominence with either intensity or duration, while its f0-cue (corresponding to rising pitch) is mostly displaced from the stressed syllable (cf. Zahner, Kutscheid, & Braun, 2019) and, being a rise, may take longer to execute (cf. Evans, 2015; Xu & Sun, 2002).

Our second research question concerned the extent to which accentual prominence and phrasal position would be expressed by similarly weighted and clustered phonetic properties in the two dialects, which might reflect physiological constraints (Hypothesis 2a) or shared control parameters (Hypothesis 2b). The overall similarity in the main dimensions of variation in the dataset was striking, but the details in terms of cross-dialectal differences were also intriguing. Specifically, SSBE's nuclear falls involved a tight correlation between f0 change and the intensity of the stressed vowel (Gramming et al., 1988; Hirano et al., 1969; Tilsen, 2016) and the nuclear pattern differed significantly from the post-nuclear pattern. In comparison, Glasgow's nuclear rises involved a weaker correlation of f0 change with intensity and a stronger one with unstressed vowel duration. High pitch on the unstressed syllable of Glasgow trochees seems to attract long duration, while attracting high intensity to a more limited extent than high pitch on the stressed syllable of SSBE trochees does. These patterns confirm what the prominence gradients suggested: that Glasgow has a rather weak concentration of prominence cues on the lexically stressed syllable itself and some of these cues are delayed beyond the stressed syllable.

It remains unclear whether the correlations of pitch and intensity, and pitch and duration, are physiologically governed or rather due to a shared control parameter as proposed by Tilsen (2016): the difference in strength of correlation points to at least tentative support for a linguistic control parameter. Importantly, our measures of pitch at the midpoints of the stressed and unstressed vowels are only rough approximations of the actual pitch accent realisations in these data, and a full analysis of the dialects' intonational categories will be needed to fully get to grips with pitch patterns. We need to establish the implications of the pitch shape and excursion size are for the relationships with duration and intensity. Even so, the Glasgow results add to the converging evidence that phrase-final lengthening and accentual lengthening are not independent (as proposed by Cummins, 1999), but interact in some way (Byrd & Riggs, 2008; Katsika, 2016; Kim et al., 2017).

Our third research question concerned whether the phonetics of accentual prominence are constrained by dialect-specific segmental phonology. Hypothesis (3) predicted more lengthening of contrastively long than short sounds, with trading relations among the prominence cues investigated. The data strongly supported this prediction. In order to maintain or exaggerate segmental contrasts under accentuation, contrastively long sounds were shown to lengthen more than contrastively short sounds in both dialects. In SSBE, accentual prominence increases duration only for those vowels that are "expandable" by virtue of being tense/diphthongal and/or followed by a [+voice] segment (cf. de Jong, 1991). In Glasgow, there is a trade-off between durational and spectral correlates of accentual prominence: /e/ and /ɪ/ lengthen under accentuation but do not become more peripheral, while other vowels show the opposite pattern. Interestingly, /ɪ/ is a central vowel in Glasgow (Stuart-Smith, 1999) and peripheralising could compromise its segmental identity. /e/ is very close in F1-F2 space to /i/. The spectral behaviour of /e/ may have to be constrained in order to preserve contrast, whereas temporal expansion could enhance its contrast with /i/, which is an extremely short vowel under most circumstances (due to the SVLR-constraints, which affect /i/ but not /e/; see 1.6). Taken together, these results demonstrate how the pressures to maintain various types of paradigmatic segmental contrast within a system can have targeted effects on the expression of prominence, cf. de Jong (1995).

A few other aspects of the results stand out. First, the data contribute converging evidence as to the domains of final lengthening, and of accentual lengthening. Final lengthening in these trochaic words in both dialects begins early: there is lengthening on vowels of both the main-stressed syllable and the final one. Accentual lengthening on the other hand was reliably found only in SSBE, where it also spread rightwards from the accented to unaccented syllable (cf Turk & White, 1999 among others). Glasgow showed very limited evidence of accentual lengthening in nuclear accent positions, which makes it typologically unusual. Future work could usefully contrast phrase-medial accented and unaccented words with non-final nuclear-accented words. If accentual lengthening is found in Glasgow in this case, it would suggest that the lack of accentual lengthening on nuclear-accented final words may be due to an interaction with final lengthening (cf. Kim et al., 2017), or durational ceiling effects in phrase-final positions (Nakai et al., 2012; Rathcke & Stuart-Smith, 2016). If accentual lengthening is not found in non-final positions either, its explanation might rather relate to the extra time needed to execute a rise (Xu & Sun, 2002; see also Evans, 2015), since rises are the most widespread nuclear pattern in Glasgow (see 1.4).

There could be interesting consequences for perception when a variety has unreliable accentual lengthening and a shallow prominence gradient. Possibly, this could make the segmental qualities present in accented syllables more difficult to detect, and/or the presence of an accent itself might be harder to detect, delaying or compromising listeners' ability to perceive the semantic and pragmatic meanings associated with prominence. These are empirical questions that are worthy of future investigation, along the lines of the eye-tracking experiments by Zahner, Kember, and Braun (2017) and Zahner, Kutscheid, and Braun (2019) that investigate the role of timing of an f0 peak in lexical access. We can tentatively suggest that both the SSBE pattern (with clustered, localised cues to prominence centered around the lexically stressed syllable) and the Glasgow pattern (with cues to prominence that are more dispersed and located after the accented syllable) might have their own advantages and disadvantages for perception. In the SSBE case, the segmental content of accented syllables and the presence/location of an accent itself should be unambiguous; but a potential correlate of strongly localised prominence cues might be lack of salience of segmental information in *non*-prominent syllables, e.g. in casual speech reduction phenomena. In the Glasgow case, the lexically stressed and accented syllable might have less intrinsic salience, but impressionistically, Glasgow nuclear rises are highly salient, especially when followed by a stretch of high pitch in the form of a plateau. Hsu, Evans, and Lee (2015) argue from combined EEG and ERP data for the perceptual salience of rises, and we know that backwards effects exist in perception (e.g., Warren & Sherman, 1974): thus it is possible that a rise, once detected, attracts attention to the material at its beginning. For the future, experiments could be devised to test whether each type of pattern places distinct stresses upon the perceptual system.

The phonological implications of the findings also deserve brief comment. It was previously known that the intonational and segmental systems of SSBE and Glasgow differ (see 2.1), and the present findings can be thought of as showing that the phonetic expression of aspects of prosodic structure differs too. But another way to capture our findings could itself be structural, i.e. in terms of differences in word-level prosody. Specifically, it has long been proposed that intervocalic consonants in English trochees are ambisyllabic (e.g. Kahn, 1976; Gussenhoven, 1986). Perhaps, the greater concentration of prominence-lending properties (intensity, duration, high pitch) on the lexically unstressed syllable means that the conditions for ambisyllabicity are not met in trochees in Glasgow, so that the intervocalic consonant in a word like *cheesy* is syllabified into the onset of the second syllable. This could in turn explain why the duration gradient flattens under accentuation in SSBE (i.e., the trochee's stressed syllable is not as much longer than the unstressed syllable in the NF condition as in the PF condition): the flatter gradient could reflect a hyperarticulation of the closed syllable in (ambisyllabic) SSBE, whereas no such pattern is found for (non-ambisyllabic) Glasgow. An interpretation in terms of differences in syllabification is also supported by the fact that Glasgow vowels seem to lack the tense/lax distinction that goes hand-in-hand with syllable structure, in the sense that lax vowels must be followed by coda consonant(s), while tense vowels need not. Further investigation of the phonetic detail of the intervocalic consonants in our data may shed light on this issue. One could expect more lenited consonantal realisations for ambisyllabic consonants than for those that are syllabified into onsets. This is consistent with our acoustic and perceptual impressions, e.g. voiced fricatives seem more likely to devoice in SSBE than in Glasgow. However, the issue warrants a study of its own, and articulatory data might be relevant (cf. Oh & Byrd, 2019).

Finally, it is worth briefly commenting on the connection of the prominence gradient to a dialect's rhythm. The concept of a durational prominence gradient as embodied in the Pairwise Variability Index and elaborated by Nolan & Asu (2009) is normally thought to relate to reduction of unstressed vowels: since the seminal work of Dauer (1983), 'syllable-timed' languages or varieties are said to be those that do not permit their unstressed vowels to reduce very far. The present findings suggest that a language or variety's tendency to have short *stressed* vowels, and a reluctance to expand these under prominence, might achieve a similar effect, especially when combined with a default intonational pattern that increases pitch, intensity and duration *after* the accented syllable rather than *on* it. These considerations underscore that rhythm metrics do not capture a single unified dimension of speech (cf. Arvaniti, 2012a, 2012b, Nolan & Asu, 2009, Rathcke & Smith, 2015).

In summary, we have shown that within a language, there can be quite substantial divergence across varieties in the way accentual prominence is expressed: we were able to find a variety of English that does not appear to lengthen accented syllables, unlike standard varieties (Turk & Sawusch, 1997, Turk & White, 1999, Turk & Shattuck-Hufnagel, 2000) and the cross-linguistic norm (Fletcher, 2010). The variation we have observed emphasises the tight connection that accentual prominence has with both the segmental and intonational structure of an utterance. Future work should investigate Glasgow's pitch categories and peak alignment in more detail, to allow a more differentiated understanding of how the intonational system influences the durational and spectral expression of prominence.

Acknowledgements

This research was funded by the UK Economic and Social Research Council, ESRC First Grant RES-061-25-0302, *Timing in Accents of English.* Thanks to Cordula Klein for assistance with data labelling, to Emma Balfour Carey for assistance with data collection, and to the three anonymous reviewers.

OPTION 1 – LINE SPACING

Appendix 1: Example dialogues

Bold font indicates nuclear accent.

***cheesy,* nuclear non-final**

A: What do you want to do about the decorations?
B: Well, I thought some disco balls and glitter everywhere.
A: That sounds a bit too **cheesy** for them.
B: For *them*? They love all that tacky glitzy stuff, though!

***cheesy,* nuclear final**

A: Ugh, it was so boring after you all left.
B: Yeah, you guys were watching that zombie film, weren't you?
A: I reckoned it was kind of **cheesy**.
B: Well, it was a zombie film after all. What did you expect?

***cheesy,* post-nuclear final**

A: So, what about playing a big chord at the end of the set?
B: Nah, I think that sounds sort of cheesy.
A: I don't think it's **remotely** cheesy.
B: Oh come on! Every Scottish country dance band ends with a big chord like that! We can't do that!

OPTION 2 – INDENTATION

Appendix 1: Example dialogues

Bold font indicates nuclear accent.

***cheesy,* nuclear non-final**

A: What do you want to do about the decorations?
B: Well, I thought some disco balls and glitter everywhere.
A: That sounds a bit too **cheesy** for them.
B: For *them*? They love all that tacky glitzy stuff, though!

***cheesy,* nuclear final**

A: Ugh, it was so boring after you all left.
B: Yeah, you guys were watching that zombie film, weren't you?
A: I reckoned it was kind of **cheesy**.
B: Well, it was a zombie film after all. What did you expect?

***cheesy,* post-nuclear final**

A: So, what about playing a big chord at the end of the set?
B: Nah, I think that sounds sort of cheesy.
A: I don't think it's **remotely** cheesy.
B: Oh come on! Every Scottish country dance band ends with a big chord like that! We can't do that!

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wocn.2019.100934.

References

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh: Edinburgh University Press.
Abercrombie, D. (1979). The accents of standard English in Scotland. In A. J. Aitken & T. McArthur (Eds.), *Languages of Scotland* (pp. 68–84). Edinburgh: Chambers.
Agutter, A. (1988). The not-so-Scottish Vowel Length Rule. In J. M. Anderson & N. MacLeod (Eds.), *Edinburgh studies in the English language* (pp. 120–132). Edinburgh: John Donald.
Aitken, A. J. (1981). The Scottish Vowel Length Rule. In M. Benskin & M. L. Samuels (Eds.), *So many people languages and tongs: Philological essays in Scots and Mediaeval English presented to Angus McIntosh* (pp. 131–157). Edinburgh: Middle English Dialect Project.
Arvaniti, A. (2012a). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics, 40*, 351–373.
Arvaniti, A. (2012b). Rhythm classes and speech perception. In O. Niebuhr (Ed.), *Understanding prosody: The role of context, function and communication (pp. 75–92)*. Berlin: Walter de Gruyter.
Barry, W. J., Andreeva, B., & Steiner, I. (2007). The phonetic exponency of phrasal accentuation in French and German. *Proceedings of Interspeech 2007, Antwerp*. 1010–1013.
Baumann, S., Becker, J., Grice, M., & Mücke, D. (2007). Tonal and articulatory marking of focus in German. In J. Trouvain & W. Barry (Eds.), *Proceedings of the 16th international congress of phonetic sciences* (pp. 1029–1032).
Baumann, S., & Winter, B. (2018). What makes a word prominent? Predicting untrained German listeners' perceptual judgements. *Journal of Phonetics, 70*, 20–38.
Beckman, M. E., & Edwards, J. (1990). Lengthening and shortening and the nature of prosodic constituency. In J. Kingston & M. E. Beckman (Eds.), *Laboratory phonology I* (pp. 152–178). Cambridge: Cambridge University Press.
Berkovits, R. (1994). Durational effects in final lengthening, gapping and contrastive stress. *Language and Speech, 37*, 237–250.

Bishop, J., Kuo, G., Kim, B. (2019). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from rapid prosody transcription. *Journal of Phonetics, 75*, 113–147.

Byrd, D., & Riggs, D. (2008). Locality interactions with prominence in determining the scope of phrasal lengthening. *Journal of the International Phonetic Association, 38*, 187–202.

Byrd, D., & Saltzman, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics, 31*, 149–180.

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language, 86*, 1–42.

Cambier-Langeveld, T., & Turk, A. (1999). A cross-linguistic study of accentual lengthening: Dutch vs English. *Journal of Phonetics, 27*, 255–280.

Cho, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics, 32*, 141–176.

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics, 37*, 466–485.

Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics, 33*, 121–157.

Cole, J., Hualde, J. I., Smith, C. L., Eager, C., Mahrt, T., & Napoleão de Souza, R. (2019). Sound, structure and meaning: The bases of prominence ratings in English, French and Spanish. *Journal of Phonetics, 75*, 113–147. https://doi.org/10.1016/j.wocn.2019.05.002.

Cole, J., Mo, Y., & Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes, 25*, 1141–1177.

Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2010). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology, 1*, 425–452.

Cooper, W. E., & Danly, M. (1981). Segmental and temporal aspects of utterance-final lengthening. *Phonetica, 38*, 106–115.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge: Cambridge University Press.

Cruttenden, A. (2007). Intonational diglossia: A case study of Glasgow. *Journal of the International Phonetic Association, 37*, 257–274.

Cummins, F. (1999). Some lengthening factors in English speech combine additively at most rates. *Journal of the Acoustical Society of America, 105*, 476–480.

Dauer, R. M. (1983). Stress-timing and syllable-timing re-analysed. *Journal of Phonetics, 11*, 51–62.

de Jong, K. J. (1991). An articulatory study of consonant-induced vowel duration changes in English. *Phonetica, 48*, 1–17.

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America, 97*, 491–504.

de Jong, K. J., & Zawaydeh, B. A. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics, 30*, 53–75.

Eefting, W. (1991). The effect of "information value" and "accentuation" on the duration of Dutch words, syllables, and segments. *Journal of the Acoustical Society of America, 89*, 412–424.

Evans, J. P. (2015). High is not just the opposite of Low. *Journal of Phonetics, 51*, 1–5.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. Hardcastle, J. Laver, & F. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 523–602). Oxford: Wiley-Blackwell.

Gramming, P., Sundberg, J., Ternström, S., Leanderson, R., & Perkins, W. (1988). Relationship between changes in voice pitch and loudness. *Journal of Voice, 2*, 118–126.

Grau, J. A., & Kemler-Nelson, D. G. (1988). The distinction between integral and separable dimensions: Evidence for the integrality of pitch and loudness. *Journal of Experimental Psychology: General, 117*, 347–370.

Gussenhoven, C. (1986). English plosive allophones and ambisyllabicity. *Gramma, 10*, 119–141.

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. Speech Prosody 2002. In *Proceedings of the First International Conference on Speech Prosody. Aix-en-Provence, ProSig and Université de Provence Laboratoire Parole et Langage* (pp. 47–57).

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, C. (2015). *Does phonological prominence exist?* (pp. 7–24). XIV: Lingue e Linguaggio.

Gussenhoven, C. (2011). Sentential prominence in English. In van Oostendorp, M., Ewen, C.J., Hume, E. & Rice, K. (eds.), The Blackwell companion to phonology. 5 vols. Malden, MA & Oxford: Wiley-Blackwell. 2778-2806.

Heldner, M., & Strangert, E. (2001). Temporal effects of focus in Swedish. *Journal of Phonetics, 29*, 329–361.

Hirano, M., Ohala, J., & Vennard, W. (1969). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *Journal of Speech, Language, and Hearing Research, 12*(3), 616.

House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America, 25*, 105–113.

Hsu, C.-H., Evans, J. P., & Lee, C.-Y. (2015). Brain responses to spoken F0 changes: Is H special? *Journal of Phonetics, 51*, 82–92.

Kahn, D. (1976). *Syllable-based generalizations in English phonology*. New York: Garland.

Katsika, A. (2016). The role of prominence in determining the scope of boundary lengthening in Greek. *Journal of Phonetics, 55*, 149–181.

Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., & Goldstein, L. (2014). The coordination of boundary tones and their interaction with prominence. *Journal of Phonetics, 44*, 62–82.

Katz, J. (2012). Compression effects in English. *Journal of Phonetics, 40*(3), 390–402.

Keating, P. A. (1985). Universal phonetics and the organization of grammars. In V. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 115–132). Orlando: Academic Press.

Kim, S., Jang, J., & Cho, T. (2017). Articulatory characteristics of preboundary lengthening in interaction with prominence on tri-syllabic words in American English. *Journal of the Acoustical Society of America, 142*, EL362–EL368.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in connected discourse. *Journal of Phonetics, 3*, 129–140.

Klatt, D. H. (1979). Synthesis by rule of segmental durations in English sentences. In B. Lindblom & S. Ohman (Eds.), *Frontiers of Speech Communication Research* (pp. 287–299). Academic Press.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America, 118*, 1038–1054.

Ladd, D. R. (2005). Vowel length in Scottish English: New data based on the alignment of pitch accent peaks. *Handout from talk presented at the 13th Manchester Phonology Meeting*.

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America, 32*, 451–454.

Lindblom, B. (1968). Temporal organization of syllable production. *Speech Transmission Quarterly Progress Status Report, 2–3*, 1–5.

Mayo, C., Aylett, M., & Ladd, D. R. (1997). Prosodic transcription of Glasgow English: an evaluation of GlaToBI. In A. Botinis, G. Kouroupetroglou, & G. Carayiannis (Eds.), *Intonation: Theory, Models and Applications (Proceedings of an ESCA Workshop, Athens, September, 1997)* (pp. 231–234). Athens: : European Speech Science Association.

McKenna, G. E. (1988). *Vowel duration in the Standard English of Scotland* Unpublished M. Litt. dissertation. University of Edinburgh.

Melara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception and Psychophysics, 482*, 169–178.

Munhall, K., Fowler, C., Hawkins, S., & Saltzman, E. (1992). "Compensatory Shortening" in monosyllables of spoken English. *Journal of Phonetics, 20*, 225–239.

Myers, S., & Hansen, B. B. (2007). The origin of vowel length neutralization in final position: Evidence from Finnish speakers. *Natural Language and Linguistic Theory, 25*, 157–193.

Nakai, S., Turk, A. E., Suomi, K., Granlund, S., Ylitalo, R., & Kunnari, S. (2012). Quantity constraints on the temporal implementation of phrasal prosody in Northern Finnish. *Journal of Phonetics, 40*, 796–807.

Nolan, F. J., & Asu, E. (2009). The Pairwise Variability Index and coexisting rhythms in language. *Phonetica, 66*, 64–77.

Oh, M. Byrd, D. (2019). Syllable-internal corrective focus in Korean. *Journal of Phonetics, 77*, 100933.

Oliver, D., & Grice, M. (2003). Phonetics and phonology of lexical stress in Polish verbs. *Proc. International Congress of Phonetic Sciences Barcelona*, 2027–2030.

Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America, 32*, 693–703.

Rathcke, T. (2017). How truncating are 'truncating languages'? Evidence from Russian and German. *Phonetica, 73*, 194–228.

Rathcke, T., & Smith, R. (2011). Exploring timing in accents of British English. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1666–1669).

Rathcke, T., & Smith, R. (2015). Speech timing and linguistic rhythm: On the acoustic bases of rhythm typologies. *JASA, 137*, 2834–2845.

Rathcke, T., & Stuart-Smith, J. (2016). On the tail of the Scottish Vowel Length Rule. *Language and Speech, 59*, 404–430.

Remijsen, B., & Gilley, L. (2008). Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *Journal of Phonetics, 36*, 318–344.

Rietveld, A. C. M., & Gussenhoven, C. (1985). On the relation between pitch excursion size and pitch prominence. *Journal of Phonetics, 13*, 299–308.

Rosenberg, A., & Hirschberg, J. (2006). On the correlation between energy and pitch accent in read English speech. *Proceedings of Interspeech – Ninth international conference on spoken language processing, Pittsburgh, PA*.

Scobbie, J. M., Hewlett, N., & Turk, A. (1999). Standard English in Edinburgh and Glasgow: The Scottish Vowel Length Rule revisited. In P. Foulkes & G. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 230–245). London: Hodder.

Scobbie, J. M., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In P. Avery, B. E. Dresher, & K. Rice (Eds.), *Contrast in phonology: Theory, perception, acquisition* (pp. 87–113). Berlin: Mouton de Gruyter.

Smith, C.L., Erickson, E., Savariaux, C. (2019). Articulatory and acoustic correlates of prominence in French: Comparing L1 and L2 speakers. *Journal of Phonetics, 77*, 100938.

Stuart-Smith, J. (1999). Glasgow: Accent and voice quality. In P. Foulkes & G. Docherty (Eds.), *Urban voices: Accent studies in the British Isles* (pp. 203–222). London: Arnold.

Stuart-Smith, J. (2003). The phonology of modern urban Scots. In J. Corbett, J. D. McClure, & J. Stuart-Smith (Eds.), *The Edinburgh companion to Scots* (pp. 110–137). Edinburgh: Edinburgh University Press.

Stuart-Smith, J. (2004). Scottish English: Phonology. In B. Kortmann, K. Burridge, E. W. Schneider, R. Mesthrie, & C. Upton (Eds.), *A Handbook of Varieties of English: 1: Phonology* (pp. 47–67). Berlin: Mouton de Gruyter.

Sullivan, J. (2011). *Approaching intonational distance and change* Unpublished Ph.D. thesis. University of Edinburgh.

Tabain, M. (2003). Effects of prosodic boundary on /aC/ sequences: Articulatory results. *Journal of the Acoustical Society of America, 113*, 2834–2849.

Terken, J. (1991). Fundamental frequency and perceived prominence of accented syllables. *Journal of the Acoustical Society of America, 89*, 1768–1776.

Terken, J. (1994). Fundamental frequency and perceived prominence of accented syllables. II non-final accents. *Journal of the Acoustical Society of America, 95*, 3662–3665.

Tilsen, S. (2016). A shared control parameter for F0 and intensity. *Proceedings of Speech Prosody, 2016*, 1066–1070.

Turk, A. E., Nakai, & Sugahara, M. (2006). Acoustic segment durations in prosodic research: A practical guide. In Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., … Schließer, J. (Eds.), Methods in empirical prosody research (pp. 1–28). Berlin: De: Gruyter.

Turk, A. E., & Sawusch, J. R. (1996). The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America, 99*, 3782–3790.

Turk, A. E., & Sawusch, J. (1997). The domain of accentual lengthening in American English. *Journal of Phonetics, 25*, 25–41.

Turk, A., & Shattuck-Hufnagel, S. (2000). Word-boundary-related duration patterns in English. Journal of Phonetics, 28(4), 397–440.

Turk, A., & Shattuck-Hufnagel, S. (2007). Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics, 35*, 445–472.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics, 27*, 171–206.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America, 58*, 434–445.

Vayra, M., & Fowler, C. A. (1992). Declination of supralaryngeal gestures in spoken Italian. *Phonetica, 49*, 48–60.

Wagner et al. (2015). Different parts of the same elephant: A roadmap to disentangle and connect different perspectives on prosodic prominence. In The Scottish Consortium for ICPhS 2015 (Ed.), Proceedings of the 18th international congress of phonetic sciences. Glasgow, UK: the University of Glasgow. ISBN 978-0-85261-941-4. Paper number 0202: 1-5, retrieved from http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0202.pdf.

Warren, R. M., & Sherman, G. L. (1974). Phonemic restorations based on subsequent context. *Perception and Psychophysics, 16*, 150–156.

Wells, J. C. (1982). *Accents of English* (Vol. 1–3) Cambridge: Cambridge University Press.

White, L.S., & Mády, K. (2008). The long and the short and the final: Phonological vowel length and prosodic timing in Hungarian. In Proceedings of the 4th international conference on speech prosody (pp. 363–366). Campinas, Brazil.

White, L., Payne, E., & Mattys, S. L. (2009). Rhythmic and prosodic contrast in Venetan and Sicilian Italian. In M. C. Vigário, S. Frota, & M. J. Freitas (Eds.), *Phonetics and phonology: Interactions and interrelations* (pp. 137–158). Amsterdam: John Benjamins.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America, 91*(3), 1707–1717.

Xu, Y., & Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America, 111*, 1399–1413.

Zahner, K., Kember, H., & Braun, B. (2017). Mind the peak: When museum is temporarily understood as musical in Australian English. *Proceedings of Interspeech*.

Zahner, K., Kutscheid, S., & Braun, B. (2019). Alignment of f0 peak in different pitch accent types affects perception of metrical stress. *Journal of Phonetics, 74*, 75–95. https://doi.org/10.1016/j.wocn.2019.02.004.